

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:45:56 ; Search time 3959 Seconds

(without alignments)  
16047.651 Million cell updates/sec

Title: US-09-745-506-74

Perfect score: 1553

Sequence: 1 GTGATGTCTTACTTCTGTCGCT.....TCTCTTACTTACATTCACAA 1553

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_com:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_un:\*

28: em\_vl:\*

29: em\_htg\_hum:\*

30: em\_htg\_inv:\*

31: em\_htg\_mus:\*

32: em\_htg\_other:\*

33: em\_htg\_pln:\*

34: em\_htg\_rtd:\*

35: em\_htg\_vit:\*

36: em\_htg\_vit:\*

37: em\_htg\_vit:\*

38: em\_htg\_vit:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1529	98.5	1606	9	AB038949 Homo sapi
2	1515	97.6	1574	6	BD158389 Primer fo
3	1515	97.6	1574	9	AK023378 Homo sapi
4	1487.6	95.8	1579	9	AF182416 Homo sapi
5	1338.4	86.2	1387	9	AF060513 Homo sapi
6	1336.6	86.1	1440	9	BC007654 Homo sapi
7	1334	85.9	1425	9	AF283538 Homo sapi
8	1232	79.3	1353	9	HS080522 Homo sapi
9	1053	67.8	1053	6	AX119075 Sequence
10	868	55.9	1836	10	AF284439 Mus muscu
11	725	46.7	796	6	BD149184 Homo sapi
12	708.4	45.6	55374	9	AL645474 Human DNA
13	417	26.9	170586	2	AC037455 Homo sapi
14	417	26.9	190508	9	AC005037 Homo sapi
15	414.4	26.7	196250	2	AC093681 Homo sapi
16	329.6	21.2	149819	2	AC121091 Mus muscu
17	329.6	21.2	234976	2	AC118698 Mus muscu
18	325.4	21.0	231600	2	AC130779 Rattus no
19	224.8	14.5	249	6	BD049005 Sequence
20	223.2	14.4	1328	3	AK114307 Clona Int
21	145.2	9.3	281017	2	AC123462 Rattus no
22	141	9.1	198250	2	AC093681 Homo sapi
23	134.6	8.7	155127	2	AL929495 Danio rer
24	117.6	7.6	231600	2	AC130779 Rattus no
25	108.4	7.0	1036	3	AY069556 Drosophill
26	108.4	7.0	70667	2	AC017306 Drosophill
27	108.4	7.0	165839	3	AC092298 Drosophill
28	108.4	7.0	273414	3	AE003650 Drosophill
29	108.4	7.0	299935	3	AE003415 Drosophill
30	97.8	6.3	157118	9	AC069286 Homo sapi
31	90	5.8	339	6	BD040994 Sequence
32	53.2	3.4	12320	1	AE013129 Thermomana
33	53	3.4	5463	6	AR227150 Sequence
34	53	3.4	28324	8	SPCC126
35	49.6	3.2	218470	1	BSUB0013 S.pombe c
36	49.6	3.2	282700	1	BACJH642 D84432 Bacillus su
37	49.6	3.2	299550	1	AP001511 Bacillus su
38	47.6	3.1	805	6	AX434353 Sequence
39	46.6	3.0	1187	8	SCYGL221C 272743 S.cerevisia
40	45.8	2.9	348250	1	AP003592 Nostoc sp
41	45.6	2.9	302173	1	AE016951 Enterococ
42	43.6	2.8	300800	1	SC0939112 Streptomy
43	43.4	2.8	209604	2	AC101593 Mus muscu
44	43.2	2.8	2000	6	AX655393 Sequence
45	43.2	2.8	299050	1	BX251411 Tropherym

## ALIGNMENTS

RESULT 1

AB038949

LOCUS Homo sapiens ALS2C1 mRNA for amyotrophic lateral sclerosis 2,

DEFINITION candidate 1, complete cds.

ACCESSION AB038949

VERSION AB038949.1 GI:12862477

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1 (sites) Hadano,S., Yanagisawa,Y., Skaug,J., Fichter,K., Nasir,J., Martindale,D., Koop,B.F., Scherer,S.W., Nicholson,D.W.,



Db 1480 AATGTCCTATCCCTCGTAGGTAAGTAACCTGTAATATACTACATATTATAAACAATG 1539  
Qy 1501 TTCTATTAACTAGAGAAAGATTGAATTAATCTGTTCTTAACATTGAA 1553  
Db 1540 TTCTATTAACTAGAGAAAGATTGAATTAATCTGTTCTTAACATTGAA 1592

RESULT 2  
BD158389  
LOCUS BD158389 1574 bp DNA linear PAT 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD158389  
VERSION BD158389.1 GI:27864147  
KEYWORDS JP 2002191363-A/13232.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1574)  
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 13232 09-JUL-2002;  
HELIX RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
PN JP 2002191363-A/13232  
PF 09-JUL-2002  
PI 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
10', C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH key  
Location/Qualifiers  
(271)..(1320).  
FEATURES  
SOURCE CDS Location/Qualifiers  
1..1574  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 420 a 361 c 372 g 421 t  
ORIGIN  
Query Match 97.6%; Score 1515; DB 6; Length 1574;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1548; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

Db 327 TGAGAGTTGGACAATGTTGGATTACTGTGGAAACAAGCCACCACTACTGTAATAC 386  
Qy 361 ACTCTTCTGACCAATGACCTGACTGAGGAAGTATGAGAGAGTGTCTGCAAAAAGAGGC 420  
Db 387 ACTCTTCTGACCAATGACCTGACTGAGGAAGTATGAGAGAGTGTCTGCAAAAAGAGGC 446  
Qy 421 AGACCTCATCTCTCCATACATCCGCGCATATCTCCGACCCATGAAAGGCATTAACCTGGAA 480  
Db 447 AGACCTCATCTCTCTCCATACATCCGCGCATATCTCCGACCCATGAAAGGCATTAACCTGGAA 506  
Qy 481 CACATGGAAGAGGCGCTGATGCCGCTCTGGAGAACAAGTGGTATCTACTCTCC 540  
Db 507 CACATGGAAGAGGCGCTGATGCCGCTCTGGAGAACAAGTGGTATCTACTCTCC 566  
Qy 541 TCATACAGCCATATATGTTGGGCCCCAGAGGGGTCAACAACCTGTGGCTAAAGGCTTGG 600  
Db 567 TCATACAGCCATATATGTTGGGCCCCAGAGGGGTCAACAACCTGTGGCTAAAGGCTTGG 626  
Qy 601 AGCTTGTACCTCCAGGCGCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGAA 660  
Db 627 AGCTTGTACCTCCAGGCGCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGAA 686  
Qy 661 CCACCGAGTAAATTAACGTTAACTACACCCCAAGACCTGGACAATGATGTGACGT 720  
Db 687 CCACCGAGTAAATTAACGTTAACTACACCCCAAGACCTGGACAATGATGTGACGT 746  
Qy 721 GAAGGAATGACGCTGTTCTGTCACCTCTTTCTGCTAGAGCTGGTAAATGAGGAACA 780  
Db 747 GAAGGAATGACGCTGTTCTGTCACCTCTTTCTGCTAGAGCTGGTAAATGAGGAACA 806  
Qy 781 AACACGATTAATCTGAATTTACTACAGAGGCTTTGATGAGAGTGTAGATTTTCTTTC 840  
Db 807 AACACGATTAATCTGAATTTACTACAGAGGCTTTGATGAGAGTGTAGATTTTCTTTC 866  
Qy 841 CCGGACAACAACACTTTATACAGAGAGCGAAATCTGTCTACTGAGAAAGCTTCTTCT 900  
Db 867 CCGGACAACAACACTTTATACAGAGAGCGAAATCTGTCTACTGAGAAAGCTTCTTCT 926  
Qy 901 ACATACGTGAATGGAGGCTTATGCACACTGGATGAATCTGTCCTCCGGAACCATGAT 960  
Db 927 ACATACGTGAATGGAGGCTTATGCACACTGGATGAATCTGTCCTCCGGAACCATGAT 986  
Qy 961 TGATCGAATTAAGAGACACACTAAACATATCTATATTCCTTGAAGCCTTGGGTGGAG 1020  
Db 987 TGATCGAATTAAGAGAGACACACTAAACATATCTATATTCCTTGAAGCCTTGGGTGGAG 1046  
Qy 1021 AACCTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGCTGCTGGTCTGGAGCAGCGTCT 1080  
Db 1047 AACCTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGCTGCTGGTCTGGAGCAGCGTCT 1106  
Qy 1081 GCAGGGTGTAGAGCTGACCTTTTACTCTACAGGTGAGATGTCCTCATATGATCTTGGGA 1140  
Db 1107 GCAGGGTGTAGAGCTGACCTTTTACTCTACAGGTGAGATGTCCTCATATGATCTTGGGA 1166  
Qy 1141 TGTGCTTCCCAAGGAATTAATGTCATCTCTGTGTAACACAGCACTGAACGAGGCTT 1200  
Db 1167 TGTGCTTCCCAAGGAATTAATGTCATCTCTGTGTAACACAGCACTGAACGAGGCTT 1226  
Qy 1201 TCTTCTGACCTTGGAGATATGCTGATCTCACTTGGAGATAAGATAATATATCTCT 1260  
Db 1227 TCTTCTGACCTTGGAGATATGCTGATCTCACTTGGAGATAAGATAATATATCTCT 1286  
Qy 1261 ATCAGAGACTACAGAGGACCTCTTCAGGTGTATATGCAAGAAATCAGAGATACAC 1320  
Db 1287 ATCAGAGACTACAGAGGACCTCTTCAGGTGTATATGCAAGAAATCAGAGATACAC 1346  
Qy 1321 ATTCCTTAACATCAAGCTGATGCCCAACTTAATTTGTAACATGATGATGAGGAGCTGT 1380  
Db 1347 ATT-CTACAAATTCAGCTGATG-CCAACTTAATTTGTAACATGATGATGAGGAGCTGT 1404  
Qy 1381 GTGCTTCCAGAGAGTGTCTTGGAGGATATCATATTCCTGCTTGTATATCTATTCACC 1440  
Db 1405 GTGCTTCCAGAGAGTGTCTTGGAGGATATCATATTCCTGCTTGTATATCTATTCACC 1464

QY 1441 AATGTCATGCTGCTGAAGTAAACCTGATATATACCATATATTAATTAACAAATG 1500  
|||||  
Db 1465 AATGTCATGCTGCTGAAGTAAACCTGATATATACCATATTAATTAACAAATG 1524  
|||||  
QY 1501 TTGATTAATTAAGTCTAGGAAGATTTGAATTAATCTGTTTACTTAACATTC 1551  
|||||  
Db 1525 TTGATTAATTAAGTCTAGGAAGATTTGAATTAATCTGTTTACTTAACATTC 1574  
|||||

RESULT 3  
AK023378 1574 bp mRNA linear PRI 01-AUG-2002  
LOCUS DEFINITION Homo sapiens CDNA FLJ13316 fls, clone OVARC1001555, weakly similar  
to NGGI-INTERACTING FACTOR 3.  
ACCESSION AK023378  
VERSION AK023378.1 GI:10435289  
KEYWORDS oligo capping, fls (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
Nakamura, Y., Nagahara, K., Masuno, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 1574)  
AUTHORS Isogai, T. and Otsuki, T.  
TITLE Direct Submission  
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,  
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan: cDNA full insert  
sequencing: Research Association for Biotechnology: cDNA library  
construction, 5'- & 3'-end one pass sequencing and clone selection:  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
FEATURES  
source Location/Qualifiers  
1. 1574  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="OVARC1001555"  
/tissue\_type="ovary, tumor tissue"  
/clone\_id="OVARC1"  
/note="cloning vector: PME18SFL3"  
271..1323  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB14551.1"  
/db\_xref="GI:10435290"  
/translation="MDIKALISLNDFAISLFAESMDNGLVPEPSPHTVNTFLTN  
DLTEVMEVLOKADKADILSYHPPIFRPKRITWNTWKRLVIRALENVGYSPTA  
YDAPOGVYNNMLAKIGACTSRPIHPSKPNPTBGNHVEYEVNVTQDLKYMSAYK  
GIDGVSVTSFSAKNGEQTIRNLCTQALMQVDFLSRNQLQKTEILSEKFL  
LATGGRICLTLDSEVSLAMIDIRIKRLKLSHRIALIGVGRTELSQKVAACAGSGS  
SVLQGVADLTLDSEVSLAMIDIRIKRLKLSHRIALIGVGRTELSQKVAACAGSGS  
INITSETDRDRLQVY"

BASE COUNT 420 a 361 c 372 g 421 t  
ORIGIN

Query Match 97.6%; Score 1515; DB 9; Length 1574;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1548; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 GTGATTTGATATCTTTGGTGTGACAGAGACAGCAAGAGAGATTGGGTACAGAAATCTGC 60  
|||||  
Db 27 GTGATTTGATATCTTTGGTGTGACAGAGACAGCAAGAGAGATTGGGTACAGAAATCTGC 86  
|||||  
QY 61 CCTGCCGACACAGACAGACAGCGCATGATGGGACAGGGGTCTGACTCAGACTTAACCTGG 120  
|||||  
Db 87 CCGCGCGACACAGAGACAGCGCATGATGGGACAGGGGTCTGACTCAGACTTAACCTGG 146  
|||||  
QY 121 CTGTGTCTGTGATTTTTCACATGCTCCCTGGAAAAAGCGCTGAAGTGGACAGAAATGAGGA 180  
|||||  
Db 147 CTGTGTCTGTGATTTTTCACATGCTCCCTGGAAAAAGCGCTGAAGTGGACAGAAATGAGGA 206  
|||||  
QY 181 TAGATGATGCCCGACAGACAGTCGCTTGTAGATTCCTGATCTGCAATTCCTCCGTTTC 240  
|||||  
Db 207 TAGATGATGCCCGACAGACAGTCGCTTGTAGATTCCTGATCTGCAATTCCTCCGTTTC 266  
|||||  
QY 241 CTTCATGATTTGAAGGCT 300  
|||||  
Db 267 CTTCATGATTTGAAGGCT 326  
|||||  
QY 301 TGAGATTTGGGACAAATGTTGATTTACTGGTGGAAACCAAGCCACACATATCTGTAATAC 360  
|||||  
Db 327 TGAGATTTGGGACAAATGTTGATTTACTGGTGGAAACCAAGCCACACATATCTGTAATAC 386  
|||||  
QY 361 ACTCTTCTGACCAATGACCTGACTGAGAGATGATGAGAGAGTCTGCAAAAGAGCC 420  
|||||  
Db 387 ACTCTTCTGACCAATGACCTGACTGAGAGATGATGAGAGAGTCTGCAAAAGAGCC 446  
|||||  
QY 421 AGACCCATTTCTCTCTACCATCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
|||||  
Db 447 AGACCCATTTCTCTCTACCATCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 506  
|||||  
QY 481 CACATGGAAGAGAGCGCTGATCCGGGCTGAGAAACAGATCGGATCTACTCTCC 540  
|||||  
Db 507 CACATGGAAGAGAGCGCTGATCCGGGCTGAGAAACAGATCGGATCTACTCTCC 566  
|||||  
QY 541 TCAATAGGCTATGATGCTGCGCCGCCAGAGGCGTCAACAACACTGGTGGCTAAAGGCTTGG 600  
|||||  
Db 567 TCAATAGGCTATGATGCTGCGCCGCCAGAGGCGTCAACAACACTGGTGGCTAAAGGCTTGG 626  
|||||  
QY 601 AGCTTATCTCTCAGAGCCCATATCATCTTCCCAAGCTCCCACTACCTACAGAGGAA 660  
|||||  
Db 627 AGCTTATCTCTCAGAGCCCATATCATCTTCCCAAGCTCCCACTACCTACAGAGGAA 686  
|||||  
QY 661 CCACGAGTAGAATTCAGATTACATACACACCAAGACCTGGACAAAGTATGTCAGT 720  
|||||  
Db 687 CCACGAGTAGAATTCAGATTACATACACCAAGACCTGGACAAAGTATGTCAGT 746  
|||||  
QY 721 GAAAGGAATTCAGAGGCTTCT 780  
|||||  
Db 747 GAAAGGAATTCAGAGGCTTCT 806  
|||||  
QY 781 AACACGATTAATCTGAATTTGATCAGAGAGGCTTGTGATGAGATTTTCTTC 840  
|||||  
Db 807 AACACGATTAATCTGAATTTGATCAGAGAGGCTTGTGATGAGATTTTCTTC 866  
|||||  
QY 841 CCGGAACAAACAACTTATCAGAGAGGAAATTCGTCTACCTGAGAGAGCTTGTCTCT 900  
|||||  
Db 867 CCGGAACAAACAACTTATCAGAGAGGAAATTCGTCTACCTGAGAGAGCTTGTCTCT 926  
|||||  
QY 901 AACTATGGAAGAGAGGCTTATGACACACTGATGTAATGTCTCCCTGGCAACATGAT 960  
|||||  
Db 927 AACTATGGAAGAGAGGCTTATGACACACTGATGTAATGTCTCCCTGGCAACATGAT 986  
|||||  
QY 961 TGATTCGAATTAAGACACCTAAACATATCTCATATTCCTTGGGCTGGGGAG 1020  
|||||  
Db 987 TGATTCGAATTAAGACACCTAAACATATCTCATATTCCTTGGGCTGGGGAG 1046  
|||||  
QY 1021 AACCTTAGAGTCAAGTCAAAAGTCTGGCCCTGTGTCTGTCTGTGGAGCAGCGCTTC 1080  
|||||  
Db 1047 AACCTTAGAGTCAAGTCAAAAGTCTGGCCCTGTGTCTGTGTGGAGCAGCGCTTC 1106  
|||||  
QY 1081 GCAAGGTGTGAGGCTGACCTTTACTCTACAGAGTGAGATGTCCATCATGATCTTTGCA 1140  
|||||

Dp	1107	GCAGGGCTGTAGGGCTGACCTTTACCTTCACAGGTGAGATGTGCCATGATGATCTTTGGCA	1166
Oy	1141	TGCTGCTTCCCAAGGAATAAATGTGCATCCTCTGTGAACACAGCAACTGACAGGCTT	1200
Dp	1167	TGCTGCTTCCCAAGGAATAAATGTGCATCCTCTGTGAACACAGCAACTGACAGGCTT	1226
Oy	1201	TCCTTTCGACCTTGAGATATGCTGGATCTTCACCTGGAGAAATAGATAAATATTAATCT	1260
Dp	1227	TCTTTTCGACCTTGAGATATGCTGGATCTTCACCTGGAGAAATAGATAAATATTAATCT	1286
Oy	1261	ATCAGAGACTACAGGAGACCTCTTCAGGTGGTATTAATGCAGAAACATCGAGTAACAC	1320
Dp	1287	ATCAGAGACTACAGGAGACCTCTTCAGGTGGTATTAATGCAGAAACATCGAGTAACAC	1346
Oy	1321	ATTCCCTCAAAATCAGCTGGATGGCCCACTTAAATTTGTAACATGACGACAGGGACGCT	1380
Dp	1347	ATT-CTCAAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGACGACAGGGACGCT	1404
Oy	1381	GTGCTTCAGAGAGATGTCTTCGAGGGATCATCATTCGCCGTTGGTAATCTTATTCACC	1440
Dp	1405	GTGCTTCAGAGAGATGTCTTCGAGGGATCATCATTCGCCGTTGGTAATCTTATTCACC	1464
Oy	1441	AAATGTCTCATGCTCGTAAGGTAAACGTGAATATATAACATATTAATTAACAATG	1500
Dp	1465	AAATGTCTCATGCTCGTAAGGTAAACGTGAATATATAACATATTAATTAACAATG	1524
Oy	1501	TTCAATTAAACTAGAAAGATTTGAATTAATCTGTTTACTTAAACATTC	1551
Dp	1525	TTCAATTAAAT-CTAGAAAGATTTGAATTAATCTGTTTACTTAAACATTC	1574

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
AF182416	Homo sapiens MDS015 (MDS015) mRNA, complete cds.	AF182416	AF182416.1	GI:10197631		Homo sapiens (human)	1579 bp				
						Homo sapiens					
						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
						1 (bases 1 to 1579)					
						Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.					
						Novel genes expressed in hematopoietic stem/progenitor cells from Myelodysplastic Syndromes patient					
						Unpublished					
						2 (bases 1 to 1579)					
						Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.					
						Direct Submission					
						Submitted (02-SEP-1999) Chinese National Human Genome Center at Shanghai, 351 Gu Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China					
						Location/Qualifiers					
						1..1579					
						/organism="Homo sapiens"					
						/mol_type="mRNA"					
						/db_xref="taxon:9606"					
						/cell_type="hematopoietic stem/progenitor cells"					
						/note="Isolated from a Myelodysplastic Syndrome patient"					
						1..1579					
						/gene="MDS015"					
						244..1299					
						/gene="MDS015"					
						/codon_start=1					
						/product="MDS015"					
						/protein_id="AAG14952.1"					
						/db_xref="GI:10197632"					
						/translation="MDKLKLSLNDPASLSPASMNGVILVSPPHVNTLFLTN					
						DLTEHWAEVYLQKKADILILSHPIIFPPMKRINLEHNGREPLVTRALENVGILTSPTT					
						AGIDAPGVNWMIAKGGAGCTSRPIHPSKANPVTGSHNVEETNVNTQDIDKMSAV					
						KGIDVATSFSAATGNEQIRIMLQTKALMVDVFLSRNQLQYKRTILSLERKL					

		LH1HGMRICLTLDSESVLAMPIDRIKRLKLSHRLALNGVFRTLESQVRYVALJAGAGS	
		SSVLQGEADLTLTGEMSHDDTLPLDASOGINVLIELSHSNTBERGLSLDRIMLDJSHLER	
		KINILSDTRDDPLQV"	
BASE COUNT	449 a	354 c	360 g
ORIGIN			416 t
Query Match	95.8%	Score 1487.6;	DB 9; Length 1579;
Best Local Similarity	99.4%;	Pred. No. 0;	
Matches 1546;	Conservative	0; Mismatches	4; Indels
			6; Gaps
			5;
QY	1	GGGATTTGATTCCTGGTCTGTGACAGGAGCAGACAGAGAAGAGGAGATTGGGTCCAGAAAACATGC	60
DB	1	GTGATTTGATTCCTGGTCTGTGACAGGAGCAGACAGAGAAGAGGAGATTGGGTCCAGAAAACATGC	60
QY	61	CTGCGCGACACAGACAGACAGGCACATAGTGGGACAGGGGCTCTGACTGAGACTTAACTG	120
DB	61	CCTGCGCGACACAGACAGGCACATAGTGGGACAGGGGCTCTGACTGAGACTTAACTG	120
QY	121	CTGTGTCTGTGCTTTTTCACCTGTCTCGAAAAGGCGCTGAAGTGGCACTGAAATGAGGCA	180
DB	121	CTGTGT -TCGTGCTTTTTCACCTGTCTCGAAAAGGCGCTGAAGTGGCACTGAAATGAGGCA	179
QY	181	TAGATGATCCCCACAGACAGTCCGGTTGTGATTCCTGTATCCCTGATATCCGCAATTCCTCCGTTG	240
DB	180	TAGATGATGTCCTCCACAGACAGTCCGGTTGTGATTCCTGTATCCCTGATATCCGCAATTCCTCCGTTG	239
QY	241	CTTCATGTGATTTGAAGGCTCTCTCTTCTCTCTGTAATGATCTTGGCATCCCTCTGCTTGGC	300
DB	240	CTTCATGTGATTTGAAGGCTCTCTCTTCTCTCTGTAATGATCTTGGCATCCCTCTGCTTGGC	299
QY	301	TGAGAGTTGGGCAATGTTGGATTACTGGTGGAAACCAAGCCACACATCTGTAATATAC	360
DB	300	TGAGAGTTGGGCAATGTTGGATTACTGGTGGAAACCAAGCCACACATCTGTAATATAC	359
QY	361	ACTCTTCTCGAACCATGACCTGACCTGAGAGAGTGAATGAGAGAGTGTCTGCAAAAAGAACGC	420
DB	360	ACTCTTCTCGAACCATGACCTGACCTGAGAGAGTGAATGAGAGAGTGTCTGCAAAAAGAACGC	419
QY	421	AGACCTCATTCCTCTCCATCCGCTATCTTCGAGCCCAATGAAAGGCA -TAACTCTGGA	479
DB	420	AGACCTCATTCCTCTCCATCCGCTATCTTCGAGCCCAATGAAAGGCAATTAACCTGGA	479
QY	480	ACACATGG - AAGGAGCGCTGTGTATCCGGGCTCTGGAGAACAGAGTCCGTTATCTACTC	537
DB	480	ACACATGGGAAAGGAGGCGCTGTGTATCCGGGCTCTGGAGAACAGAGTCCGTTATCTACTC	539
QY	538	TCCTCATACAGCTTATGATGCTGCGCCCGCAGGCGCTCAACAATCTGTGCTTAAAGGCT	597
DB	540	TCCTCATACAGCTTATGATGCTGCGCCCGCAGGCGCTCAACAATCTGTGCTTAAAGGCT	599
QY	598	TGGAGCTTGTACCTCCAGGCGCATTCATCTTCCAAACCTCCCAACTCCCTACAGAGG	657
DB	600	TGGAGCTTGTACCTCCAGGCGCATTCATCTTCCAAACCTCCCAACTCCCTACAGAGG	659
QY	658	AAACCAACGAGTAGAATTCAAAGCTTAACTACACCACCAAGACCTGACCAAAAGTCATGCTGCG	717
DB	660	AAACCAACGAGTAGAATTCAAAGCTTAACTACACCACCAAGACCTGACCAAAAGTCATGCTGCG	719
QY	718	AGTGAAGAAGATTGACGGTGTCTGTGCACTTCTTTTCTGCTAGAGACTGGTATGAGGA	777
DB	720	AGTGAAGAAGATTGACGGTGTCTGTGCACTTCTTTTCTGCTAGAGACTGGTATGAGGA	779
QY	778	ACAAACAGGAGTATTCGAAATTTGATTCAGAAAGGCTTGAATGCAAGGTGGTGAATTTCT	837
DB	780	ACAAACAGGAGTATTCGAAATTTGATTCAGAAAGGCTTGAATGCAAGGTGGTGAATTTCT	839
QY	838	TTCCCGGAACAAACAATTTATCAGAAAGGAAATTTGTCTACTGTGAGAAAGCTTTGCT	897
DB	840	TTCCCGGAACAAACAATTTATCAGAAAGGAAATTTGTCTACTGTGAGAAAGCTTTGCT	899
QY	898	TCTACATCTGGAATGGACGGTTATGCAACACTGGATGAATCTGTCTCCCTGGCAACAT	957
DB	900	TCTACATCTGGAATGGACGGTTATGCAACACTGGATGAATCTGTCTCCCTGGCAACAT	959

QY 958 GATTGATCGAATAAAAAGACACCTAAACATATCTATTCGCTTACGCTTGGGGTGGG 1017  
 DB 960 GATTGATCGAATAAAAAGACACCTAAACATATCTATTCGCTTACGCTTGGGGTGGG 1019  
 QY 1018 GAGAACCTTGAAGTCTCAAGTCAAAAGTGTGGCCCTGTGTGCTGTTGCGGAGCAGCGT 1077  
 DB 1020 GAGAACCTTGAAGTCTCAAGTCAAAAGTGTGGCCCTGTGTGCTGTTGCGGAGCAGCGT 1079  
 QY 1078 TCTGAGGGTGTGAGGCTGACCTTACCTCAACAGGTGAGATGCTCCATCATGATATCTT 1137  
 DB 1080 TCTGAGGGTGTGAGGCTGACCTTACCTCAACAGGTGAGATGCTCCATCATGATATCTT 1139  
 QY 1138 GGAATGCTGTCCCAAGAAATATGTCATCTCTGTGAACACAGCAACACTGAACAGAG 1197  
 DB 1140 GGAATGCTGTCCCAAGAAATATGTCATCTCTGTGAACACAGCAACACTGAACAGAG 1199  
 QY 1198 CTTTCTTCTGACCTTCGAGATATGCTGATCTCTACTTGGAGAAATTAATTAATTTAT 1257  
 DB 1200 CTTTCTTCTGACCTTCGAGATATGCTGATCTCTACTTGGAGAAATTAATTAATTTAT 1259  
 QY 1258 CCTATCAGAGACTGACAGAGGACCCCTCTCAGGTGATATGTCAGAGAAACATCAGATTA 1317  
 DB 1260 CCTATCAGAGACTGACAGAGGACCCCTCTCAGGTGATATGTCAGAGAAACATCAGATTA 1319  
 QY 1318 CACATTCCTCAAAATCAGCTGATGCGCAACTTAATTTGTAAACATGAGTCACTGGAGCT 1377  
 DB 1320 CACATTCCTCAAAATCAGCTGATGCGCAACTTAATTTGTAAACATGAGTCACTGGAGCT 1377  
 QY 1378 GGTGCTCTCCAGAGAGTGTCTTCGAGGTATCATATTCGCTTGTAACTATATC 1437  
 DB 1378 GGTGCTCTCCAGAGAGTGTCTTCGAGGTATCATATTCGCTTGTAACTATATC 1437  
 QY 1438 ACCAATGTTCTATCGCTCGTAAAGTAACTGTAATTAATTAATTAATTAATTAACAA 1497  
 DB 1438 ACCAATGTTCTATCGCTCGTAAAGTAACTGTAATTAATTAATTAATTAATTAACAA 1497  
 QY 1498 ATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1553  
 DB 1498 ATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1553  
 RESULT 5  
 AF060513 1387 bp mRNA linear PRI 02-JAN-2001  
 LOCUS AF060513 Homo sapiens clone 016d06 My018 protein mRNA, complete cds.  
 DEFINITION AF060513  
 ACCESSION AF060513  
 VERSION AF060513.1 GI:12001975  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 1387)  
 AUTHORS Mao, X.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-APR-1998) Institute of Genetics, School of Life  
 Science, Fudan University, 220 Handan Rd., Shanghai 200433,  
 P.R.China  
 FEATURES  
 source  
 1..1387  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="016d06"  
 /tissue\_type="brain"  
 /dev\_stage="fetus"  
 /dev\_start=1  
 /product="My018 protein"  
 /protein\_id="AA043131.1"  
 /db\_xref="GI:12001975"  
 /translation="MDKALISSLNDPASFSLFASRWNVGLVESPPTVNTLFLTN

BASE COUNT 378 a 323 c 309 g 377 t  
 ORIGIN  
 Query Match 86.2%; Score 1338.4; DB 9; Length 1387;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1361; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
 DLEEVMEVLOKADLLISHPPIFRPKRRTITWTKERLIVIRALENVGISPTHTA  
 YDAAPGVNNMLAKIGAGCTSRPIHPSKAPNPTBENHNEVNVTODLDKMSAVK  
 GIDGVSFSEFARNEBOTRIINLNTOKALQVDFLSRNLQVOKETIISLEKPL  
 LHMGRCLTDESVALMIDIRIKRLEKLSHLRALGIVRFLSESVKVAACAGGS  
 SYLQGVADLLYLTGEMSHDITDAASQGINVILCEHSNTEREPLSLDRMLDSHLENK  
 INITLSETRDPLQVY"  
 188 GTCCCCACGACAGCTCCGGTTTGTAGATTCCTGATCTGCAATTCCTCCGTTCTCANG 247  
 DB 25 GTCCCCACGACAGCTCCGGTTTGTAGATTCCTGATCTGCAATTCCTCCGTTCTCANG 84  
 QY 248 GATTGAAGGCTCTCTTCTTCTTCTGATGACTTGATTCATCCCTCTGCTTGTGAGACT 307  
 DB 85 GATTGAAGGCTCTCTTCTTCTTCTGATGACTTGATTCATCCCTCTGCTTGTGAGACT 144  
 QY 308 TGGGACAATGTTGATTACTGTTGGAACCAAGCCACACATCTGTAATACACTCTTC 367  
 DB 145 TGGGACAATGTTGATTACTGTTGGAACCAAGCCACACATCTGTAATACACTCTTC 204  
 QY 368 CTGACCAATGACCTGACTGAGAGAGTATGAGAGAGTGTCTGCAAAAAGAGCAGACTTC 427  
 DB 205 CTGACCAATGACCTGACTGAGAGAGTATGAGAGAGTGTCTGCAAAAAGAGCAGACTTC 264  
 QY 428 ATTCCTCTCTACATCCGCTATCTCCCAACCCATGAAGGCTAAACCTGGAAACATATG 487  
 DB 265 ATTCCTCTCTACATCCGCTATCTCCCAACCCATGAAGGCTAAACCTGGAAACATATG 324  
 QY 488 AAGAGCGCTGTGTGATCCGGGCTCTGAGAACAGAGTGTCTACTCTCTCTCATACA 547  
 DB 325 AAGAGCGCTGTGTGATCCGGGCTCTGAGAACAGAGTGTCTACTCTCTCTCATACA 384  
 QY 548 GCTTATGATGCTGCGCCCGCCAGAGGCTCAACACTGTTGGCTAAAGGCTTGGAGCTTGG 607  
 DB 385 GCTTATGATGCTGCGCCCGCCAGAGGCTCAACACTGTTGGCTAAAGGCTTGGAGCTTGG 444  
 QY 608 ACCTCCAGGCGCATATATCTTCCAAAGCTCCCAATACCTCTACAGAGGAAACACCGA 667  
 DB 445 ACCTCCAGGCGCATATATCTTCCAAAGCTCCCAATACCTCTACAGAGGAAACACCGA 504  
 QY 668 GTAGAAATCAACGTTAATCAACCAAGAGCTGGCAAAAGTGTCTGCAAGTGAAGA 727  
 DB 505 GTAGAAATCAACGTTAATCAACCAAGAGCTGGCAAAAGTGTCTGCAAGTGAAGA 564  
 QY 728 ATTGAGGGTGTCTGCTACTTCTTTTCTGCTAGAGCTGTATGAGAAACAACAGG 787  
 DB 565 ATTGAGGGTGTCTGCTACTTCTTTTCTGCTAGAGCTGTATGAGAAACAACAGG 624  
 QY 788 ATTAATCTGAATGTACTCTGAGAAGGCTTGATGAGCTGTGATATTTCTTCCGGAAC 847  
 DB 625 ATTAATCTGAATGTACTCTGAGAAGGCTTGATGAGCTGTGATATTTCTTCCGGAAC 664  
 QY 848 AAACCACTTATCAGAGAGAGGAATTTGTCTACTGAGAGAGCTTGTCTTCACTACT 907  
 DB 685 AAACCACTTATCAGAGAGAGGAATTTGTCTACTGAGAGAGCTTGTCTTCACTACT 744  
 QY 908 GGAATGGAGCGTTATGACACACTGATGAATCTGTCTCCCTGGCAACATGATGATCGA 967  
 DB 745 GGAATGGAGCGTTATGACACACTGATGAATCTGTCTCCCTGGCAACATGATGATGCA 804  
 QY 968 ATTAAGAAAGACACTAATAATCTCATATTCCTGTAGCCCTTGGGTGGGGAACACTTA 1027  
 DB 805 ATTAAGAAAGACACTAATAATCTCATATTCCTGTAGCCCTTGGGTGGGGAACACTTA 864  
 QY 1028 GAGTCTCAAGTCAAACTCGGCGCTGTGTGCTGCTTGGAGAGAGCTTGTGACAGGCT 1087  
 DB 865 GAGTCTCAAGTCAAACTCGGCGCTGTGTGCTGCTTGGAGAGAGCTTGTGACAGGCT 924





QY	787	GATTAATCTGTAATCTGACAGAGGCTTGTGATGAGGTGTAGATTTCTTTCCCGGA	846
DB	660	GATTAATCTGTAATCTGACAGAGGCTTGTGATGAGGTGTAGATTTCTTTCCCGGA	719
QY	847	CAAAACATTTTTCAGAGGAGGAATCTGTCACTGAGAGACCTTGTCTTCTACATAC	906
DB	720	CAAAACATTTTTCAGAGGAGGAATCTGTCACTGAGAGACCTTGTCTTCTACATAC	779
QY	907	TGATATGGAGCGTATGTCACACTGATGATCTGTCTCCCTGGACACCATGATGATG	966
DB	780	TGATATGGAGCGTATGTCACACTGATGATCTGTCTCCCTGGACACCATGATGATG	839
QY	967	AATAAAGACACCTTAAACCTATCTATCTGCTTACGCTTGGGGTGGAGAACCTT	1026
DB	840	AATAAAGACACCTTAAACCTATCTATCTGCTTACGCTTGGGGTGGAGAACCTT	899
QY	1027	AGAGCTCAAGTCAAGTCTGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1086
DB	900	AGAGCTCAAGTCAAGTCTGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	959
QY	1087	TGTTGAGGCTGACCTTACCTCAGAGTGTGAGATGTCCTCATGATCTTGTGATGCTC	1146
DB	960	TGTTGAGGCTGACCTTACCTCAGAGTGTGAGATGTCCTCATGATCTTGTGATGCTC	1019
QY	1147	TTCCCAAGGATTAATGTCTATCTCTGTGACACAGCAACACTGAAAGGCTTCTTTC	1206
DB	1020	TTCCCAAGGATTAATGTCTATCTCTGTGACACAGCAACACTGAAAGGCTTCTTTC	1079
QY	1207	TGACCTTGAGATGTGAGTGTGATCTGACCTGAGATTAAGATTAATTAATCTTACCA	1266
DB	1080	TGACCTTGAGATGTGAGTGTGATCTGACCTGAGATTAAGATTAATTAATCTTACCA	1139
QY	1267	GACTGACAGGAGCCCTCTCAGAGTGTGATTAATGAGAAACATCAGATACACATCTC	1326
DB	1140	GACTGACAGGAGCCCTCTCAGAGTGTGATTAATGAGAAACATCAGATACACATCTC	1198
QY	1327	ACAAATCAGCTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1386
DB	1199	ACAAATCAGCTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1257
QY	1387	CCAGAGAGTGTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1446
DB	1238	CCAGAGAGTGTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1317
QY	1447	TCTATCGCTGCTGAGTAAACCTGATTAATTAATTAATTAATTAATTAATTAATTA	1506
DB	1318	TCTATCGCTGCTGAGTAAACCTGATTAATTAATTAATTAATTAATTAATTAATTA	1377
QY	1507	ATAAATCTGAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1553
DB	1378	ATAAATCTGAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1424

RESULT 7  
AF283538 1425 bp mRNA linear PRI 16-JAN-2001  
LOCUS AF283538  
DEFINITION Homo sapiens NIF3L1 protein mRNA, complete cds.  
ACCESSION AF283538  
VERSION AF283538.1 GI:12006402  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Tascou, S., Uedelhoven, J., Dikens, C., Nayeria, K., Engel, W. and  
Burlind, P.  
TITLE Isolation and characterization of a novel human gene, NIF3L1, and  
its mouse ortholog, Nif3l1, highly conserved from bacteria to  
mammals  
JOURNAL Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)  
MEDLINE 20573864  
PUBMED 11124544

QY	2	(bases 1 to 1425)	846
DB	2	(bases 1 to 1425)	719
QY	847	CAAAACATTTTTCAGAGGAGGAATCTGTCACTGAGAGACCTTGTCTTCTACATAC	906
DB	720	CAAAACATTTTTCAGAGGAGGAATCTGTCACTGAGAGACCTTGTCTTCTACATAC	779
QY	907	TGATATGGAGCGTATGTCACACTGATGATCTGTCTCCCTGGACACCATGATGATG	966
DB	780	TGATATGGAGCGTATGTCACACTGATGATCTGTCTCCCTGGACACCATGATGATG	839
QY	967	AATAAAGACACCTTAAACCTATCTATCTGCTTACGCTTGGGGTGGAGAACCTT	1026
DB	840	AATAAAGACACCTTAAACCTATCTATCTGCTTACGCTTGGGGTGGAGAACCTT	899
QY	1027	AGAGCTCAAGTCAAGTCTGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1086
DB	900	AGAGCTCAAGTCAAGTCTGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	959
QY	1087	TGTTGAGGCTGACCTTACCTCAGAGTGTGAGATGTCCTCATGATCTTGTGATGCTC	1146
DB	960	TGTTGAGGCTGACCTTACCTCAGAGTGTGAGATGTCCTCATGATCTTGTGATGCTC	1019
QY	1147	TTCCCAAGGATTAATGTCTATCTCTGTGACACAGCAACACTGAAAGGCTTCTTTC	1206
DB	1020	TTCCCAAGGATTAATGTCTATCTCTGTGACACAGCAACACTGAAAGGCTTCTTTC	1079
QY	1207	TGACCTTGAGATGTGAGTGTGATCTGACCTGAGATTAAGATTAATTAATCTTACCA	1266
DB	1080	TGACCTTGAGATGTGAGTGTGATCTGACCTGAGATTAAGATTAATTAATCTTACCA	1139
QY	1267	GACTGACAGGAGCCCTCTCAGAGTGTGATTAATGAGAAACATCAGATACACATCTC	1326
DB	1140	GACTGACAGGAGCCCTCTCAGAGTGTGATTAATGAGAAACATCAGATACACATCTC	1198
QY	1327	ACAAATCAGCTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1386
DB	1199	ACAAATCAGCTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1257
QY	1387	CCAGAGAGTGTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1446
DB	1238	CCAGAGAGTGTGAGTGTGATTAATGATTAATGATTAATGATTAATGATTAATGAT	1317
QY	1447	TCTATCGCTGCTGAGTAAACCTGATTAATTAATTAATTAATTAATTAATTAATTA	1506
DB	1318	TCTATCGCTGCTGAGTAAACCTGATTAATTAATTAATTAATTAATTAATTAATTA	1377
QY	1507	ATAAATCTGAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1553
DB	1378	ATAAATCTGAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1424

REFERENCE 2 (bases 1 to 1425)  
AUTHORS Tascou, S., Burlind, P. and Engel, W.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUN-2000) Institute for Human Genetics, University of  
Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany  
FEATURES  
source  
1. 1425  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
CDS  
47..1160  
/codon\_start=1  
/product="NIF3L1 protein"  
/protein\_id="AAG4846.1"  
/db\_xref="GI:12006403"  
/translation="MLSCVPRPVTVFVDSIIONSRSPMDLKALLSINDPASTS  
FAESNDNGLVEPSPPHVTNLTFLNDLIEVMEVYQKRAIDLSTHPPFRPMK  
ITWNTWKEFLVRALENRGVTSPTATDAQGVNMLDAGLACTSRPHPSAPN  
YPTBNHREENVNTODIDKVMASVKGIDGVSYSFSARGENEOTRIINLCQKAL  
MOVDPFLRNKOLYKTEILSLERPLLLHTGMRICLTDESYSLATMIDRIKRLKLS  
HRLAGRTLESQVKKVVALCAGSSSVLOGEADLYTGMESHHDPLDAASQGINV  
ILCESNHERGLSPDRMDSHLEPNKINIILSEDRDPLQV"  
BASE COUNT 385 a 331 c 314 g 395 t  
ORIGIN  
Query Match 85.9%; Score 1314; DB 9; Length 1425;  
Best local Similarity 99.9%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 2; Gaps 2;





QY 1078 TCTGACGGGTGTTAGAGCTGACCTTACCTACACAGGTGAGATGTCCTCATCANGACTTT 1137  
|||||  
Db 781 TCTGACGGGTGTTAGAGCTGACCTTACCTACACAGGTGAGATGTCCTCATCANGACTTT 840  
|||||  
QY 1138 GGATGCTGCTGCCAAGAAATATGTCATCCCTCTGTGAACACAGAACATGAAAGAG 1197  
|||||  
Db 841 GGATGCTGCTGCCAAGAAATATGTCATCCCTCTGTGAACACAGAACATGAAAGAG 900  
|||||  
QY 1198 CTTTCTTCTGACCTTCGAGATATGCTGATTCCTACTTGAGAAATAGATAATATAT 1257  
|||||  
Db 901 CTTTCTTCTGACCTTCGAGATATGCTGATTCCTACTTGAGAAATAGATAATATAT 960  
|||||  
QY 1258 CCTATCAGAGACTGACAGGAGACCTCTTCAGGTGATATATGACAGAAACATCAGATTA 1317  
|||||  
Db 961 CCTATCAGAGACTGACAGGAGACCTCTTCAGGTGATATATGACAGAAACATCAGATTA 1020  
|||||  
QY 1318 CACATTCCTCAAAATAGCTGGATGCCCACTTAATTTGTAACATGACTCAGTGGACT 1377  
|||||  
Db 1021 CACATTCCTCAAAATAGCTGGATGCCCACTTAATTTGTAACATGACTCAGTGGACT 1078  
|||||  
QY 1378 GGTGCTCTCCAGAGAGTGTCTTCGAGGATATCATTTCCGTTTGTATCTTATTC 1437  
|||||  
Db 1079 GGTGCTCTCCAGAGAGTGTCTTCGAGGATATCATTTCCGTTTGTATCTTATTC 1138  
|||||  
QY 1438 ACCAAATGTTCTATGCTGCTGTAAGTAAACTGTATATATACATACATATTAATAACAA 1497  
|||||  
Db 1139 ACCAAATGTTCTATGCTGCTGTAAGTAAACTGTATATATACATACATATTAATAACAA 1198  
|||||  
QY 1498 ATGTTTCATTTAACTCTAGAAAGATTTGTAATTAATCTGTTACTTAATCA 1553  
|||||  
Db 1199 ATGTTTCATTTAACTCTAGAAAGATTTGTAATTAATCTGTTACTTAATCA 1254  
|||||  
RESULT 9  
AX119075 1053 bp DNA linear PAT 11-MAY-2001  
LOCUS AX119075  
DEFINITION Sequence 239 from Patent WO0129221.  
ACCESSION AX119075  
VERSION AX119075.1 GI:14036029  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1  
AUTHORS Konklin,D.C. and Yee,D.P.  
TITLE Proteins and polynucleotides encoding them  
JOURNAL Patent: WO 0129221-A 239 26-Apr-2001;  
Zymogenetics, Inc. (US)  
FEATURES  
source  
1..1053  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..1053  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC38611.1"  
/db\_xref="GI:14036030"  
CDS  
1..1053  
/translation="MDIKALISLIDFASLFAESMDNYGLVESPPTVTMTPLTN  
DITEVMEVLDKKADLITSTPPIPRPKRTITWTWKRLVTALENVGTISPTA  
YDAAPGVNNMLAKGIGACTSRPIHPSKAPNPTEGNNHVEVNVNTDLDKVMASV  
GIDGVSVTSFASRTGNEQTRINLNCTOKALMOVDFLSRNOLXKTEILSLEKPL  
LHTGVRICCTLDESIVLAIMDIRIKRHLKLSHRIALRGVRLTESQKVVALCAGSGS  
SVLQGEADLYLTGEMSHDITDAASQGINVITLCHESNTERGLSDLRMDLSHLENK  
INITSETDROLDLOYV"  
BASE COUNT 279 a 253 c 247 g 274 t  
ORIGIN  
Query Match 67.8%; Score 1053; DB 6; Length 1053;  
Best Local Similarity 100.0%; Pred. No.1.4e-282;  
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 ATGATTTGAAGGCTCTCTTCTTCTTCTGAAATGACTTTCATCCCTCGTTTGTGAG 304  
|||||  
Db 1 ATGATTTGAAGGCTCTCTTCTTCTTCTGAAATGACTTTCATCCCTCGTTTGTGAG 60  
|||||  
QY 305 AGTTGGACAAATTTGGATTACATGCTGAGAACCAAGCCACACATCTGTAATATACATC 364  
|||||  
Db 61 AGTTGGACAAATTTGGATTACATGCTGAGAACCAAGCCACACATCTGTAATATACATC 120  
|||||  
QY 365 TTCCCTACCAATGACTGCTGAGAGATGAGAGAGTGTGCAAAAAGAGGAGAG 424  
|||||  
Db 121 TTCCCTACCAATGACTGCTGAGAGATGAGAGAGTGTGCAAAAAGAGGAGAG 180  
|||||  
QY 425 CTCATTCCTCTCCATCCATCCGCTATCTTCGACCCATGAAAGGCACTGGAACACA 484  
|||||  
Db 181 CTCATTCCTCTCCATCCATCCGCTATCTTCGACCCATGAAAGGCACTGGAACACA 240  
|||||  
QY 485 TGGAGAGAGCCCTGATCGGGCTCTGAGAACAGAGTGGATCTACTCTCTCAT 544  
|||||  
Db 241 TGGAGAGAGCCCTGATCGGGCTCTGAGAACAGAGTGGATCTACTCTCTCAT 300  
|||||  
QY 545 ACAGCCTATGATGCTGCGGCCCGAGGGCGTCAACACTGGTTGGCTAAAGGGCTGGAGCT 604  
|||||  
Db 301 ACAGCCTATGATGCTGCGGCCCGAGGGCGTCAACACTGGTTGGCTAAAGGGCTGGAGCT 360  
|||||  
QY 605 TGTACCTCCAGGCCATACATCCCTTCCAAAGCTCCCACTACCTACAGAGGAAACAC 664  
|||||  
Db 361 TGTACCTCCAGGCCATACATCCCTTCCAAAGCTCCCACTACCTACAGAGGAAACAC 420  
|||||  
QY 665 CGAGTGAATTCACGTTTAACCTACACCCAAAGACTGGACAAAGTCACTGCTGAGTAA 724  
|||||  
Db 421 CGAGTGAATTCACGTTTAACCTACACCCAAAGACTGGACAAAGTCACTGCTGAGTAA 480  
|||||  
QY 725 GGAATGAGCGTCTTCTGCACTCTTTTCTGCTAGAGCTGTAATAGAAACAACA 784  
|||||  
Db 481 GGAATGAGCGTCTTCTGCACTCTTTTCTGCTAGAGCTGTAATAGAAACAACA 540  
|||||  
QY 785 CGGATTAATCTGAATTTGTAAGGCTTTGATGAGAGTGTGATATTTCTTCCGG 844  
|||||  
Db 541 CGGATTAATCTGAATTTGTAAGGCTTTGATGAGAGTGTGATATTTCTTCCGG 600  
|||||  
QY 845 AACAAACACTTTATAGAAAGACGAAATTTCTGCTGAGAGAGCTTGTCTTACAT 904  
|||||  
Db 601 AACAAACACTTTATAGAAAGACGAAATTTCTGCTGAGAGAGCTTGTCTTACAT 660  
|||||  
QY 905 ACTGAATGAGAGGTTATGACACATGATGATCTGCTCCCTGGCAACATGATGAT 964  
|||||  
Db 661 ACTGAATGAGAGGTTATGACACATGATGATCTGCTCCCTGGCAACATGATGAT 720  
|||||  
QY 965 CGAATAAAAAGACACTTAAACTATATCTGATTTGCTTACCCCTGGGGTGGGAGAAC 1024  
|||||  
Db 721 CGAATAAAAAGACACTTAAACTATATCTGATTTGCTTACCCCTGGGGTGGGAGAAC 780  
|||||  
QY 1025 TTGAGATCTCAATCAAACTCGAGGCCCTGTGCTGCTGCTGGAGACAGCTTCGACG 1084  
|||||  
Db 781 TTGAGATCTCAATCAAACTCGAGGCCCTGTGCTGCTGCTGGAGACAGCTTCGACG 840  
|||||  
QY 1085 GGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCATCATGATGATCTTTGATGCT 1144  
|||||  
Db 841 GGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCATCATGATGATGATCTTTGATGCT 900  
|||||  
QY 1145 GCTTCCCAAGGAATTAATGTCATCTCTGTGGAACACAGACACTGACGAGGCTTCTT 1204  
|||||  
Db 901 GCTTCCCAAGGAATTAATGTCATCTCTGTGGAACACAGACACTGACGAGGCTTCTT 960  
|||||  
QY 1205 TCTGACCTCGAGATATGCTGATCTCTCATCTTGGAGAAATTAATTAATTCATCA 1264  
|||||  
Db 961 TCTGACCTCGAGATATGCTGATCTCTCATCTTGGAGAAATTAATTAATTCATCA 1020  
|||||  
QY 1265 GAGACTGACAGGACCTCTTCAGTGTATAA 1297  
|||||  
Db 1021 GAGACTGACAGGACCTCTTCAGTGTATAA 1053  
|||||

RESULT 10  
 AF284439 1836 bp mRNA linear F0D 16-JAN-2001  
 LOCUS Mus musculus Nif3l1 mRNA, complete cds.  
 DEFINITION  
 ACCESSION AF284439.1 GI:12034693  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Tascou, S., Deleghoven, J., Dijkens, C., Neyerla, K., Engel, W. and  
 Burfeind, P.  
 TITLE  
 Isolation and characterization of a novel human gene, Nif3l1, and  
 its mouse ortholog, Nif3l1, highly conserved from bacteria to  
 mammals  
 JOURNAL  
 Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)  
 MEDLINE  
 20573864  
 PUBMED  
 11124544  
 2 (bases 1 to 1836)  
 Tascou, S., Burfeind, P. and Engel, W.  
 AUTHORS  
 Title Submission  
 Submitted (02-JUL-2000) University of Göttingen, Institute of Human  
 Genetics, Heinrich-Duker Weg 12, Göttingen 37073, Germany  
 FEATURES  
 SOURCE  
 1..1836  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="1"  
 508..1638  
 /codon\_start=1  
 /product="Nif3l1"  
 /protein\_id="AA045961.1"  
 /db\_xref="GI:12034694"  
 /translation="MSSAHIVPTSVORAGSWICRSHFMDLKALLSLNDPASTLSE  
 AESDNVGLVEPSPHTNLFLENDLEEDALOKKADFLSYHPFIRPKHI  
 TMTWKELVRALENRYAVYSPHAYAAAGVSMYAKIGCTGTRPHSRADY  
 PTEGAHREFSVNSODLDKYMSTLRGGVSVSPRACGEIDRISLNTQTLIM  
 QVLAFLSDROLQKTEFLSLKPLILHTGKRLCTLDSESLAIVIRIKTHLISH  
 LRLALGVERLESQVKVALCAGSGSVLQGEADLYLTGEMSHDVIDVAASKGINVI  
 LCHESNTERGFLSELOEMGVHEFKINIIISERDRLRV"  
 BASE COUNT 462 a 517 c 450 g 407 t  
 ORIGIN  
 Query Match 55.9%; Score 868; DB 10; Length 1836;  
 Best Local Similarity 85.4%; Pred. No. 6.9e-231;  
 Matches 967; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

Db 828 GAAGAGTGTCTGATGATCCGGGCTCTGGAGAAACAGAGTCTGTACTCTCCACAC 887  
 Qy 547 AGCTATATATGCTGGCCCGGAGGCTGCAACAACGTGGCTTAAGGCTTGAGCTTG 606  
 Db 888 AGCTATATATGCTGGCCCGGAGGCTGCAACAACGTGGCTTAAGGCTTGAGCTTG 947  
 Qy 607 TACCTCCAGGCGCATATACATCTTCCAAAGCTCCCAACCTCTTACAGAGGAAACACCG 666  
 Db 948 CACTACACAGGCCCATTCACACCTCTTCCAGAGCTCCAGACTACCCACAGAGGAGCTCACCG 1007  
 Qy 667 ACTAGAAATTCAGTAACTACACCCAGACCTTGACAAAGTATGTCTGACGTGAAGG 726  
 Db 1008 ACTAGAAATTCAGTAACTACACCCAGACCTTGACAAAGTATGTCTGACGTGAAGG 1067  
 Qy 727 AATTCAGGCTGTTCTGTCTACTCTTTTCTCTGAGACTGGTATGAGAACCAACAGC 786  
 Db 1068 GGTGTGAGGTGTCTGTCTACTCTTTTCTCTGAGACTGGTATGAGAACCAACCG 1127  
 Qy 787 GATTAATCTGATTTACTCAGAGGCTTGTATGACAGGTGAGATTTCTTCCCGGAA 846  
 Db 1128 GATTCAGGCTGATTTACTCAGAGGCTTGTATGACAGGTGAGATTTCTTCCCGGAA 1187  
 Qy 847 CAACAACTTTATCAGAGGCGAAATTTCTGCTACTGAGAAACCTTTGCTTCAATAC 906  
 Db 1188 CAGACAACTTTATCAGAGGCGAAATTTCTGCTACTGAGAAACCTTTGCTTCAATAC 1247  
 Qy 907 TGGAAATGGAGGCTTATGACACCTGATGAATCTGTCTCCCGGCAACATATGATG 966  
 Db 1248 TGGAAATGGAGGCTTATGACACCTGATGAATCTGTCTCCCGGCAATATATGAGCG 1307  
 Qy 967 AATTAAGAGACACTTAAATCTATCTATCTATGCTTACCTTGGGCTGGGAGAACCTT 1026  
 Db 1308 AATCAAAACACACTTAAATCTATCTATCTATGCTTACCTTGGGCTGGGAGAACCTT 1367  
 Qy 1368 AAGAGTCCCAAGTCAAAAGTGTGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1427  
 Db 1087 TGTTCAGGCTGACCTTATCTACACAGGTGAGATGTCATGATGATGATGATGATG 1146  
 Qy 1428 AGTGGAGGCGGACCTTATCTACACAGGTGAGATGTCATGATGATGATGATGATG 1487  
 Db 1147 TTCCCAAGAAATTAATGTCATCTCTGTCAGACAGCAACACTGAGAGGCTTCTTTC 1206  
 Qy 1488 TTCCCAAGGATTAATGTCATCTCTGTCAGACAGCAACACTGAGAGGCTTCTTTC 1547  
 Db 1207 TGACCTTGCAGATATGCTGATGATTCATCTTGAAGATAATATATATCTATCAGA 1266  
 Qy 1548 TGACCTTGCAGATATGCTGATGATTCATCTTGAAGATAATATATATCTATCAGA 1607  
 Db 1267 GACTGACAGGAGGACCTTCTGAGGTGATATTCAGAAACATCAGAGATAC 1318  
 Qy 1608 GACAGACAGGAGGACCTTCTGAGGTGATATTCAGAAACATCAGAGATAC 1659  
 Db

RESULT 11  
 BD149184 796 bp DNA linear PAT 17-JAN-2003  
 LOCUS  
 DEFINITION  
 ACCESSION BD149184.1 GI:27854942  
 VERSION  
 KEYWORDS JP 2002191363-A/4027.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 Ota, T., Isogai, T., Nishikawa, T., Haysht, K., Saito, K., Yamamoto, J.,  
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
 TITLE  
 Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)

	PN	JP 2002191363-A/4027	
	PD	09-JUL-2002	
	PF	28-JUL-2000 JP 2000280990	
	PI	TOSHIO OTA,TAKAO ISOGAI,'TETSUO NISHIKAWA,KOJI HAYASHI,KAORU SAITO,	
	P.I.	JUNICHI YAMAMOTO,SHIZUKO ISHII,'TOMOYASU SUGIYAMA,AI WAKAMATSU,	
	PI	KEIICHI NAGAI,'TETSUJI OTSUKI	
	PC	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ 10, PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key	
FEATURES			
SOURCE	FT	source	1..796
		Location/Qualifiers	/Organism='Homo sapiens (human)'
			1..796
			/Organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'
BASE COUNT		192 a	204 c 202 g 195 t 3 others
ORIGIN			

Query Match	46.7%	Score 725	DB 6	Length 796
Best Local Similarity	98.4%	Pred. No. 5.4e-191		
Matches	762	Conservative	0	Mismatches 8; Indels 4; Gaps 3
QY	1	GTGATTTGTTATCTTGGTGTCTGCAGAGGACAGCAGAAAGAGGATTTGGTTCAGAAACTGCG	60	
Db	27	GTGATTTGTTATCTTGGTGTCTGCAGAGGACAGCAGAAAGAGATTTGGTTCAGAAACTGCG	86	
OY	61	CCTGCGGACACAGAGCAGAGGCGCTATGTGGACAGCGGGGTCTGTACTCAGACTTAACTGCG	120	
Db	87	CCTGCGGACACAGAGCAGAGGCGCTATGTGGACAGCGGGGTCTGTACTCAGACTTAACTGCG	146	
OY	121	CTGTGTCTCTGTGTTTTCACCTGTCTCTGAAAAGGCGCTGAAGTGGCAGCTGAATGAGGCA	180	
Db	147	CTGTGTCTCTGTGTTTTCACCTGTCTCTGAAAAGGCGCTGAAGTGGCAGCTGAATGAGGCA	206	
OY	181	TAGATGATATCCCCACAGACAGATCGGGTTGTAGATTCCCTGATCTCGAATTCCTCCGCTTC	240	
Db	207	TAGATGATATCCCCACAGACAGATCGGGTTGTAGATTCCCTGATCTCGAATTCCTCCGCTTC	266	
OY	241	CTTCATGATTTGGAAGGCTCTCCTTCTTCCTTGAATACCTTTGGATCCCTCTCGTTGCG	300	
Db	267	CTTCATGATTTGGAAGGCTCTCCTTCTTCCTTGAATACCTTTGGATCCCTCTCGTTGCG	326	
OY	301	TGAGAGTTGGGACAAATGTTGGATTACTGTGTGAGACCAAGCCACACATCTGTAATATAC	360	
Db	327	TGAGAGTTGGGACAAATGTTGGATTACTGTGTGAGACCAAGCCACACATCTGTAATATAC	386	
OY	361	ACTCTTCTTGACCAATGACACTGACAGAGAAAGTATGTAGAGAGGTGCTGCAAAAGAGGC	420	
Db	387	ACTCTTCTTGACCAATGACACTGACAGAGAAAGTATGTAGAGAGGTGCTGCAAAAGAGGC	446	
OY	421	AGACCTCAATCTCTCTCAACATCCGCGCTATCTTCCGACCACATGAAGCGCATTAACCTGGA	480	
Db	447	AGACCTCAATCTCTCTCAACATCCGCGCTATCTTCCGACCACATGAAGCGCATTAACCTGGA	506	
OY	481	CACATGGAAGAGCGCCTGTGTGATCCGGGCTCTGGAAGACAGATCGGTATCTACTCTCC	540	
Db	507	CACATGGAAGAGCGCCTGTGTGATCCGGGCTCTGGAAGACAGATCGGTATCTACTCTCC	566	
OY	541	TCATTCACAGCTATGATGTGCGGCGCCCAAGGCGCTCAACAACATGGTTGGCTAAAGGGCTTGG	600	
Db	567	TCATTCACAGCTATGATGTGCGGCGCCCAAGGCGCTCAACAACATGGTTGGCTAAAGGGCTTGG	626	
OY	601	AGCTTGTACTCTCCAGGCCCATACATCTTCCAAAGCTCCCACTACCTACAGAGGGA	660	
Db	627	AGCTTGTACTCTCCAGGCCCATACATCTTCCAAAGCT -CCAACCTACCTACAGAGGGA	685	
OY	661	CCACCGAATGAATTCAAAGTTAACTATACACCAAGAACTGGACAAAGTCATCTTGCAGT	720	

[illegible]

COMMENT

On Feb 1, 2002 this sequence version replaced q1:18135182.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations. Where differences are found these are annotated as variations. Together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP. Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP5-925F19 is from the library RPEC1-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP5-925F19 it may be shorter because we sequence overlapping sections only once, except for a short overlap. The true left end of clone RP11-478P15 is at 53375 in this sequence. The true right end of clone RP11-393M18 is at 2000 in this sequence.

FEATURES

SOURCE

Location/Qualifiers

1..55374

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="1"

/clone="RP5-925F19"

/clone\_1lb="RPEC1-5"

/clone\_1lb="RP5-925F19"

BASE COUNT 17426 a 9900 c 9927 g 18121 t

ORIGIN

Query Match 45.6%; Score 708.4; DB 9; Length 55374;

Best Local Similarity 76.0%; Pred. No. 4.1e-186;

Matches 1046; Conservative 0; Mismatches 231; Indels 99; Gaps 10;



```
* 3794 3893: gap of unknown length
* 3894 5414: contig of 1521 bp in length
* 5415 5514: gap of unknown length
* 5515 8773: contig of 3259 bp in length
* 8774 8874 10904: gap of unknown length
* 10905 11005 11004: gap of unknown length
* 13205 13305 13204: contig of 2200 bp in length
* 13305 16401: gap of unknown length
* 16402 16501: contig of 3097 bp in length
* 16502 18227: gap of unknown length
* 18228 18327: contig of 1726 bp in length
* 18328 22401: gap of unknown length
* 22402 22501: contig of 4074 bp in length
* 22502 25177: gap of unknown length
* 25178 25277: contig of 2676 bp in length
* 25278 29892: gap of unknown length
* 29893 33457: gap of unknown length
* 33458 33557: contig of 3465 bp in length
* 33558 37405: gap of unknown length
* 37406 37505: contig of 3848 bp in length
* 37506 41734: gap of unknown length
* 41735 41834: contig of 4229 bp in length
* 41835 48989: gap of unknown length
* 48990 49099: contig of 7165 bp in length
* 49100 56076: gap of unknown length
* 56077 56176: contig of 6977 bp in length
* 56177 63604: gap of unknown length
* 63605 63704: contig of 7428 bp in length
* 63705 70821: gap of unknown length
* 70822 70921: contig of 7117 bp in length
* 70922 80138: gap of unknown length
* 80139 80238: contig of 9217 bp in length
* 80239 88780: gap of unknown length
* 88781 88880: contig of 8542 bp in length
* 88881 104642: gap of unknown length
* 104643 104742: contig of 15762 bp in length
* 104743 124656: gap of unknown length
* 124657 124756: contig of 19914 bp in length
* 124757 143891: gap of unknown length
* 143892 143991: contig of 19135 bp in length
* 143992 170586: gap of unknown length
* 170586 170586: contig of 26595 bp in length.

FEATURES
    source
        1. 170586
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="02"
            /clone="RP11-422L5"
            /clone_1lb="RP11-11"
            1. 11171
                /note="assembly_name:Contig1"
            1272. 2319
                /note="assembly_name:Contig13"
            2420. 3793
                /note="assembly_name:Contig17"
            3894. 5414
                /note="assembly_name:Contig18"
            5515. 8773
                /note="assembly_name:Contig19"
            8874. 10904
                /note="assembly_name:Contig20"
            clone_end: 17
            11005. 13204
                /note="assembly_name:Contig21"
            13305. 16401
                /note="assembly_name:Contig22"
            16502. 18227
                /note="assembly_name:Contig23"
            18328. 22401
                /note="assembly_name:Contig24"
```

```
misc_feature 22502..25177
              /note="assembly_name:Contig25"
misc_feature 25278..29892
              /note="assembly_name:Contig26"
misc_feature 29993..33457
              /note="assembly_name:Contig27"
misc_feature 33558..37405
              /note="assembly_name:Contig28"
misc_feature 37506..41734
              /note="assembly_name:Contig29"
misc_feature 41835..48989
              /note="assembly_name:Contig30"
misc_feature 49100..56076
              /note="assembly_name:Contig31"
              clone_end: 56
              56177..63604
              /note="assembly_name:Contig32"
misc_feature 63705..70821
              /note="assembly_name:Contig33"
misc_feature 70922..80138
              /note="assembly_name:Contig34"
misc_feature 80239..88780
              /note="assembly_name:Contig35"
misc_feature 88881..104642
              /note="assembly_name:Contig36"
misc_feature 104743..124656
              /note="assembly_name:Contig37"
misc_feature 124757..143891
              /note="assembly_name:Contig38"
misc_feature 143992..170586
              /note="assembly_name:Contig39"

BASE COUNT 50417 a 34173 c 33290 g 50259 t 2447 others
ORIGIN
```

Query Match 26.9%; Score 417; DB 2; Length 170586;  
Best Local Similarity 99.5%; Pred. No. 1.2e-104;  
Matches 439; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
QY 1111 AGGTGAGATGCCATCATGATATCTTGAGTGTCTCCCAAGATAATGTCATCCT 1170
DB 128601 AGGTGAGATGCCATCATGATATCTTGAGTGTCTCCCAAGATAATGTCATCCT 128542

QY 1171 CTGTGAACACAGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
DB 128541 CTGTGAACACAGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 128482

QY 1231 TCACCTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1290
DB 128481 TCACCTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 128422

QY 1291 GGTATTAATTCGAAACATCAGATTAACACATTCCTACAAATCAGCTGGATGCCAAT 1350
DB 128421 GGTATTAATTCGAAACATCAGATTAACACATTCCTACAAATCAGCTGGATGCCAAT 128364

QY 1351 AATTTGTAACTAGATCAGTGGAGTGGTGTCTTCCAGAGATGTCTTCCAGAGATTC 1410
DB 128363 AATTTGTAACTAGATCAGTGGAGTGGTGTCTTCCAGAGATGTCTTCCAGAGATTC 128304

QY 1411 ATCATTCGCGTGTGTTAATCTTAATCACCAGATTTCTATCCTGCTGAAGTAAACTG 1470
DB 128303 ATCATTCGCGTGTGTTAATCTTAATCACCAGATTTCTATCCTGCTGAAGTAAACTG 128244

QY 1471 TAATTAATCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1530
DB 128243 TAATTAATCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128184

QY 1531 AATTCGTTACTTAACATTC 1551
DB 128183 AATTCGTTACTTAACATTC 128163

RESULT 14
AC005037/c
```





```

repeat_region 12190..12334
/rpt_family="Alu"
repeat_region 12335..12360
/rpt_family="L1"
repeat_region 12464..12514
/rpt_family="L1"
repeat_region 12515..12822
/rpt_family="Alu"
repeat_region 12823..12862
/rpt_family="L1"
repeat_region 12857..13096
/rpt_family="L1"
repeat_region 13101..13330
/rpt_family="MAlR"
repeat_region 13331..13634
/rpt_family="Alu"
repeat_region 13635..13692
/rpt_family="MAlR"
repeat_region 13704..13990
/rpt_family="L1"
repeat_region 14082..14387
/rpt_family="Alu"
repeat_region 14402..14704
/rpt_family="Alu"
repeat_region 14797..15095
/rpt_family="Alu"
repeat_region 15162..15196
/rpt_family="Alu"
repeat_region 15606..15681
/rpt_family="Mariner"
repeat_region 15691..15827
/rpt_family="L1"
repeat_region 15861..16189
/rpt_family="MER1_type"
repeat_region 16426..16446
/rpt_family="L2"
repeat_region 16447..16730
/rpt_family="(CAAT)n"
repeat_region 17072..17179
/rpt_family="Alu"
repeat_region 17212..17504
/rpt_family="L2"
repeat_region 17869..18139
/rpt_family="Alu"
repeat_region 18168..18464
/rpt_family="Alu"
repeat_region 18606..19155
/rpt_family="Alu"
repeat_region 19454..19767
/rpt_family="MAlR"
repeat_region 19807..19949
/rpt_family="Alu"
repeat_region 19950..20247
/rpt_family="MER1_type"
/rpt_family="Alu"

Query Match 26.98; Score 417; DB 9; Length 190508;
Best Local Similarity 99.5%; Pred.No.1.2e-104;
Matches 439; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

```

Db 110814 GGTATATTTGCAAGAAACATCAGATATACACATT-CTACAAATCAGCTGATG-CCACATT 110757
OY 1351 AATTGTATACATGAGTCACTGAGTGGTGTCTTCAGAGAGTCTTCGAGGTATC 1410
Db 110756 AATTGTATACATGAGTCACTGAGTGGTGTCTTCAGAGAGTCTTCGAGGTATC 110697
OY 1411 ATCATTTCCGGTTTGTATATCTTATTCACCAATGTTCTATGCTGTGAAGTAAACTG 1470
Db 110696 ATCATTTCCGGTTTGTATATCTTATTCACCAATGTTCTATGCTGTGAAGTAAACTG 110637
OY 1471 TAATATACATACATTTTAAATACAAATGTTCTATTCATTAACCTAGGAAGATTGATA 1530
Db 110636 TAATATACATACATTTTAAATACAAATGTTCTATTCATTAACCTAGGAAGATTGATA 110577
OY 1531 AATCTGTTTACTTAACTATTC 1551
Db 110576 AATCTGTTTACTTAACTATTC 110556

RESULT 15
AC093681 198250 bp DNA linear HTG 09-OCT-2001
LOCUS AC093681
DEFINITION Homo sapiens chromosome 4 clone RP11-663N2, *** SEQUENCING IN
ACCESSION AC093681.2 GI:15987253
VERSION AC093681.2
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 198250)
AUTHORS Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
TITLE Unpublished
AUTHORS 2 (bases 1 to 198250)
JOURNAL Waterston,R.H.
REFERENCE Direct Submission
AUTHORS Submitted (07-SEP-2001) Genome Sequencing Center, Washington
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Oct 9, 2001 this sequence version replaced gi:15487524.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
Project Information
Center project name: H.NH063N02
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1912: contig of 1912 bp in length
1913 2012: gap of unknown length
2013 3539: contig of 1527 bp in length
3540 3639: gap of unknown length
3640 6127: contig of 2488 bp in length
6128 6227: gap of unknown length
6228 9244: contig of 3017 bp in length
9245 9345: gap of unknown length
9345 13288: contig of 3944 bp in length
13289 13389: gap of unknown length
13389 16301: contig of 2913 bp in length
16302 16401: gap of unknown length
16401 20096: contig of 3655 bp in length
20097 20196: gap of unknown length

```



**THIS PAGE BLANK (USPTO)**